Courbe des taux JGB
Masterclass ACP

Des fondements mathématiques de l'analyse en composantes principales
aux stratégies butterfly neutres aux facteurs

Données : rendements à maturité constante de la JSDA (`jgb_cm.csv`)

Ténors : $\{2, 5, 7, 10, 15, 20, 25, 30, 35, 40\}$ ans

Nombre de facteurs par défaut : 3 (Niveau / Pente / Courbure)

Table des matières

1 Introduction **1**

2 Fondements mathématiques de l'ACP **3**
2.1 Définition mathématique . . . 3
2.1.1 Matrice de covariance . . . 3
2.1.2 Décomposition en valeurs propres . . . 3
2.1.3 Équivalence avec la SVD . . . 3
2.1.4 Ratio de variance expliquée . . . 4
2.1.5 Scores des composantes principales et reconstruction . . . 4
2.2 Vérification avec des données synthétiques . . . 4

3 Préparation des données **7**
3.1 Données de rendements à maturité constante de la JSDA . . . 7
3.2 Sélection des ténors . . . 7

4 Ingénierie des variables **9**
4.1 Variations quotidiennes des rendements . . . 9
4.2 Distribution des variations quotidiennes . . . 9
4.3 Structure de corrélation . . . 10

5 Application de l'ACP à la courbe JGB **11**
5.1 Ajustement de l'ACP . . . 11
5.2 Vérification du ratio de variance expliquée . . . 11
5.3 Scree plot . . . 12
5.4 Série temporelle des scores de facteurs . . . 12

6 Interprétation économique **15**
6.1 Normalisation du signe . . . 15
6.2 Forme des chargements . . . 15
6.2.1 PC1 : Niveau (Level) . . . 15
6.2.2 PC2 : Pente (Slope) . . . 16
6.2.3 PC3 : Courbure (Curvature) . . . 16
6.3 Décomposition de la contribution des facteurs (reconstruction) . . . 16

7 De l'ACP au trading **19**
7.1 Transactions butterfly neutres aux facteurs . . . 19
7.2 Solution de norme minimale par SVD . . . 19

8 Backtest walk-forward **21**
8.1 Principes de conception . . . 21
8.2 Mécanisme d'élimination du biais . . . 21
8.3 Logique de trading . . . 22

8.4 Traitement des coûts de transaction 22
8.5 Indicateurs de performance 22
8.6 Résultats du backtest 22

9 Validation de la robustesse 25
9.1 Sensibilité à la fenêtre d'apprentissage 25
9.2 Sensibilité aux seuils Z 25
9.3 Sensibilité au nombre de facteurs 26

10 Estimation de la fenêtre optimale et stabilité des composantes 27
10.1 Estimation de la fenêtre optimale 27
10.1.1 Compromis biais–variance 27
10.1.2 Erreur de reconstruction OOS 27
10.1.3 Stabilité de l'EVR 28
10.2 Variation temporelle des chargements 29
10.2.1 Évaluation quantitative par similarité cosinus 29
10.2.2 Carte de chaleur de la variation temporelle des chargements 29
10.3 Préservation de l'interprétabilité et mélange des composantes 30
10.3.1 Basculement des PC 30
10.3.2 Contre-mesures pratiques 30

11 Mises en garde et extensions 33
11.1 Mises en garde 33
11.2 Sujets avancés 33

Liste des symboles 35

Chapitre 1

Introduction

Ce document poursuit les trois objectifs suivants :

1. Décomposer les variations quotidiennes de la courbe JGB à maturité constante en **composantes principales orthogonales**.
2. Attribuer une **interprétation économique** aux facteurs obtenus (Niveau / Pente / Courbure).
3. Construire une stratégie butterfly neutre aux facteurs et évaluer son efficacité au moyen d'une **validation walk-forward hors échantillon (OOS)**.

Point clé

Saisir la variation de la courbe des taux dans un espace de faible dimension est une tâche centrale sur les marchés obligataires — que ce soit pour la gestion des risques, la construction de portefeuille ou le trading de valeur relative. L'analyse en composantes principales (ACP) est largement utilisée comme la méthode standard et puissante à cette fin.

Chapitre 2

Fondements mathématiques de l'ACP

Dans ce chapitre, avant d'appliquer l'ACP à des données réelles, nous vérifions sa définition mathématique à l'aide de données synthétiques bidimensionnelles.

2.1 Définition mathématique

Définition 2.1 (Analyse en composantes principales)**.** Étant donné une matrice de données centrée $\mathbf{X} \in \mathbb{R}^{n \times p}$ (n : nombre d'échantillons, p : nombre de variables), l'analyse en composantes principales est définie par la procédure suivante.

2.1.1 Matrice de covariance

La matrice de covariance empirique est définie par

$$\mathbf{\Sigma} = \frac{1}{n-1} \mathbf{X}^\top \mathbf{X}. \tag{2.1}$$

$\mathbf{\Sigma}$ est une matrice réelle symétrique semi-définie positive de taille $p \times p$.

2.1.2 Décomposition en valeurs propres

La décomposition en valeurs propres de $\mathbf{\Sigma}$ s'écrit

$$\mathbf{\Sigma} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^\top, \qquad \mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p), \qquad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0. \tag{2.2}$$

Ici $\mathbf{V} = [\, \boldsymbol{v}_1 \mid \boldsymbol{v}_2 \mid \cdots \mid \boldsymbol{v}_p \,]$ est une matrice orthogonale, et chaque vecteur colonne $\boldsymbol{v}_k$ est appelé **k-ième direction principale** (chargement, *loading*).

2.1.3 Équivalence avec la SVD

En écrivant la décomposition en valeurs singulières (SVD) de $\mathbf{X}$ sous la forme

$$\mathbf{X} = \mathbf{U} \mathbf{S} \mathbf{V}^\top, \tag{2.3}$$

les directions principales coïncident avec les colonnes de $\mathbf{V}$, et la relation entre les valeurs propres et les valeurs singulières est donnée par

$$\lambda_k = \frac{s_k^2}{n-1}, \tag{2.4}$$

où s_k est le k-ième élément diagonal (valeur singulière) de $\mathbf{S}$.

Remarque 2.1. La SVD calcule les composantes principales sans former explicitement la matrice de covariance, et est donc numériquement stable. L'implémentation interne de `sklearn.decomposition` repose également sur la SVD.

2.1.4 Ratio de variance expliquée

Le **ratio de variance expliquée** (EVR) de la k-ième composante principale est défini par

$$\mathrm{EVR}_k = \frac{\lambda_k}{\sum_{j=1}^{p} \lambda_j} = \frac{\lambda_k}{\mathrm{tr}(\boldsymbol{\Sigma})}. \tag{2.5}$$

L'identité $\sum_{k=1}^{p} \mathrm{EVR}_k = 1$ est vérifiée.

2.1.5 Scores des composantes principales et reconstruction

La **matrice des scores** obtenue en projetant la matrice de données $\mathbf{X}$ sur les directions principales est

$$\boldsymbol{F} = \mathbf{X}\mathbf{V} \in \mathbb{R}^{n \times p}. \tag{2.6}$$

La transformation inverse (reconstruction) est

$$\mathbf{X} = \boldsymbol{F}\mathbf{V}^{\top}, \tag{2.7}$$

qui restitue exactement les données d'origine. La reconstruction approchée n'utilisant que les $r < p$ premières composantes est

$$\hat{\mathbf{X}}_r = \boldsymbol{F}_{:,1:r}\, \mathbf{V}_{:,1:r}^{\top}, \tag{2.8}$$

qui, d'après le théorème d'Eckart–Young, fournit la meilleure approximation de rang r au sens de la norme de Frobenius.

2.2 Vérification avec des données synthétiques

Pour confirmer notre compréhension théorique, nous générons des données à partir d'une loi normale bidimensionnelle corrélée.

Exemple 2.1 (Données synthétiques bidimensionnelles)**.** Soit la matrice de covariance vraie

$$\Sigma_{\text{vraie}} = \begin{pmatrix} 3.0 & 2.5 \\ 2.5 & 4.0 \end{pmatrix},$$

et générons $n = 300$ échantillons.

Code 2.1 – Génération et centrage des données synthétiques

```
np.random.seed(42)
n_toy = 300
cov_true = [[3.0, 2.5], [2.5, 4.0]]
X_toy = np.random.multivariate_normal([0, 0], cov_true, n_toy)
X_centered = X_toy - X_toy.mean(axis=0)
```

Nous effectuons la décomposition en valeurs propres de la matrice de covariance empirique et la comparons aux valeurs propres obtenues par SVD :

Code 2.2 – Vérification de l'équivalence entre décomposition en valeurs propres et SVD

```
cov_sample = np.cov(X_centered, rowvar=False)
eigenvalues, eigenvectors = np.linalg.eigh(cov_sample)
idx = np.argsort(eigenvalues)[::-1]
eigenvalues, eigenvectors = eigenvalues[idx], eigenvectors[:, idx]

U, S, Vt = np.linalg.svd(X_centered, full_matrices=False)
eigenvalues_from_svd = S**2 / (n_toy - 1)
```

La figure 2.1 montre les directions principales et le scree plot pour les données synthétiques. Dans le panneau de gauche, on peut confirmer que PC1 correspond à la direction de plus grande variance ($\approx 82\%$) et PC2 à la direction orthogonale ($\approx 18\%$).

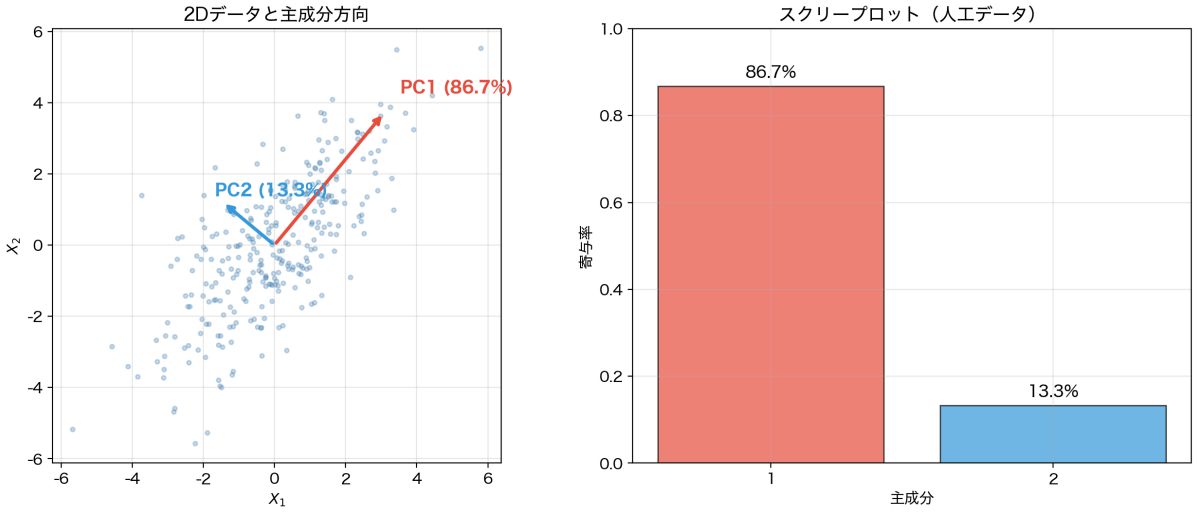

FIGURE 2.1 – Directions principales (à gauche) et scree plot (à droite) des données synthétiques 2D.

Dans la reconstruction utilisant toutes les composantes principales, l'erreur maximale par rapport aux données d'origine est de l'ordre de $O(10^{-14})$, confirmant la restitution exacte prédite par la théorie.

Point clé

- La décomposition en valeurs propres de la matrice de covariance et la SVD donnent les **mêmes directions principales** (au signe près, arbitraire).
- La reconstruction utilisant toutes les composantes coïncide avec les données d'origine à la précision machine près.
- La somme des ratios de variance expliquée vaut exactement 1.

Chapitre 3

Préparation des données

3.1 Données de rendements à maturité constante de la JSDA

Nous utilisons les données de rendements de référence à maturité constante (Constant-Maturity) des JGB de gré à gré publiées par la Japan Securities Dealers Association (JSDA). Le fichier de données `jgb_cm.csv` contient les rendements (en %) pour plusieurs ténors à chaque jour ouvré.

Code 3.1 – Chargement des données

```
raw = pd.read_csv('data/datasets/jgb_cm.csv', encoding='utf-8-sig')
raw['datetime'] = pd.to_datetime(raw['datetime'])
rate_cols = [c for c in raw.columns if c.endswith('y_rate')]
```

3.2 Sélection des ténors

Nous sélectionnons les 10 ténors suivants pour l'analyse :

$$\mathcal{T} = \{2, 5, 7, 10, 15, 20, 25, 30, 35, 40\} \quad \text{(ans)}. \tag{3.1}$$

Cela fournit une coupe transversale de la courbe des taux de dimension $p = 10$. La figure 3.1 présente des instantanés de la courbe des taux sur la période des données.

Attention

Le problème de roll du CTD dans la zone 7 ans : le titre CTD (Cheapest-to-Deliver) des contrats à terme JGB se situe souvent dans la zone 7 ans. Lors des rollovers de contrats à terme et des changements de titre CTD, des sauts peuvent apparaître dans les rendements à maturité constante. Ces sauts se manifestent comme des valeurs aberrantes dans les variations quotidiennes et peuvent affecter les résultats de l'ACP. En particulier, autour des rollovers de contrats, des discontinuités peuvent être introduites dans les variations de rendement de la zone 7 ans, ce qui nécessite une attention particulière.

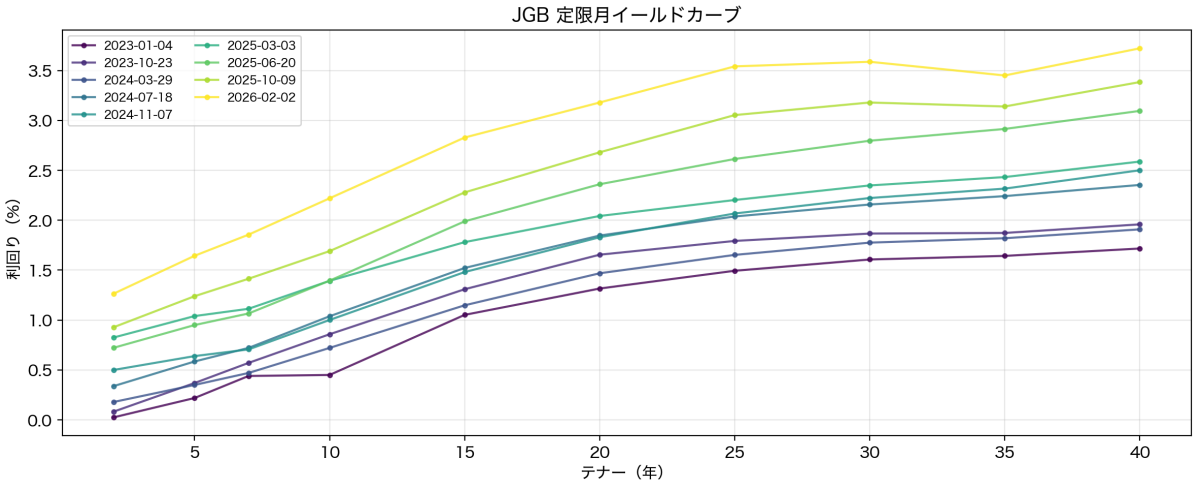

FIGURE 3.1 – Évolution de la courbe des taux JGB à maturité constante.

Chapitre 4

Ingénierie des variables

4.1 Variations quotidiennes des rendements

Comme entrée de l'ACP, nous utilisons les **variations quotidiennes** (en bps) plutôt que les niveaux de rendement. La variation quotidienne du ténor k est calculée par

$$\Delta y_t^{(k)} = (y_t^{(k)} - y_{t-1}^{(k)}) \times 100 \quad \text{(bps)}, \tag{4.1}$$

où $y_t^{(k)}$ est le rendement (en %) du ténor k à la date t.

Remarque 4.1 (Pourquoi les variations plutôt que les niveaux)**.** Si les niveaux de rendement sont injectés directement dans l'ACP, PC1 capture quasiment le niveau lui-même (la composante de tendance) et la structure de co-mouvement quotidien ne peut pas être extraite. En passant aux variations quotidiennes, la stationnarité est assurée et le motif de co-mouvement entre ténors devient clair.

Code 4.1 – Construction de la matrice des variations quotidiennes

```
def build_d1_bps(df, tenors, min_obs=60):
    cols = [f'{t}y' for t in tenors]
    d1 = df[cols].diff() * 100    # % -> bps
    d1 = d1.dropna()
    assert len(d1) >= min_obs
    return d1
```

On obtient ainsi une matrice $n \times 10$ $\mathbf{X}$ dont les lignes sont les dates et les colonnes les ténors. Cette matrice $\mathbf{X}$ constitue l'entrée de toute l'analyse ACP qui suit.

4.2 Distribution des variations quotidiennes

La figure 4.1 montre la distribution des variations quotidiennes des rendements pour chaque ténor. Les formes sont globalement proches d'une loi normale, mais pour de nombreux ténors on observe un excès de kurtosis positif (queues épaisses). En particulier, dans la zone des maturités courtes, les mouvements discontinus accompagnant les changements de politique de la Banque du Japon tendent à apparaître comme des valeurs aberrantes.

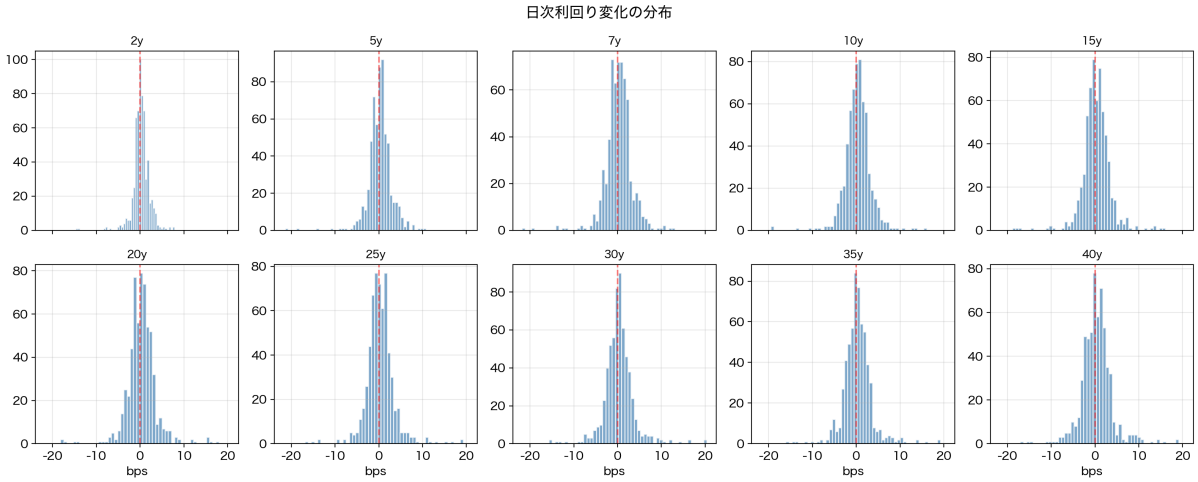

FIGURE 4.1 – Distribution des variations quotidiennes des rendements pour chaque ténor.

4.3 Structure de corrélation

La figure 4.2 montre la matrice de corrélation des variations quotidiennes. On observe les motifs typiques suivants :

— Les **ténors adjacents** présentent une forte corrélation positive ($\rho > 0.8$).

— Les **ténors éloignés** (par exemple 2 ans vs. 40 ans) présentent une corrélation plus faible.

— Cette structure de corrélation « en bande » est précisément la base sur laquelle l'ACP peut comprimer les données en un petit nombre de facteurs.

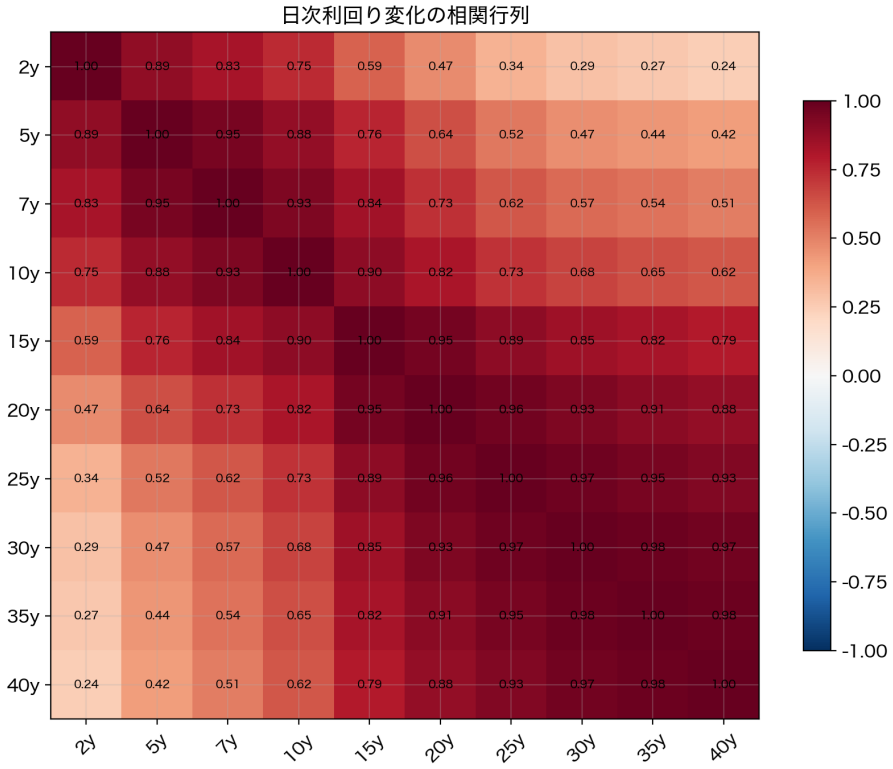

FIGURE 4.2 – Matrice de corrélation des variations quotidiennes des rendements.

Chapitre 5

Application de l'ACP à la courbe JGB

5.1 Ajustement de l'ACP

Pour la matrice des variations quotidiennes $\mathbf{X} \in \mathbb{R}^{n\times 10}$, nous exécutons une ACP avec les $r = 3$ premières composantes.

Code 5.1 – Fonction d'ajustement de l'ACP

```
def fit_pca(X, n_factors=3, standardize=False):
    X_arr = X.values.copy()
    if standardize:
        scaler = StandardScaler()
        X_arr = scaler.fit_transform(X_arr)
    pca = PCA(n_components=n_factors)
    scores_arr = pca.fit_transform(X_arr)
    loadings = pd.DataFrame(pca.components_,
                            index=[f'PC{i+1}' for i in range(
                               n_factors)],
                            columns=X.columns)
    scores = pd.DataFrame(scores_arr, index=X.index,
                          columns=[f'PC{i+1}' for i in range(
                             n_factors)])
    return {'loadings': loadings, 'scores': scores,
            'evr': pca.explained_variance_ratio_,
            'eigenvalues': pca.explained_variance_, 'pca': pca}
```

Remarque 5.1 (Faut-il standardiser ?)**.** Dans cette analyse, nous fixons `standardize=False` et exécutons l'ACP sur les variations de rendement en bps (sur une base de covariance). En conséquence, les ténors à plus forte volatilité (la zone des maturités longues) reçoivent naturellement un poids plus important. En effet, l'« échelle en bps » des variations de rendement est directement liée au risque réel, ce qui constitue un choix raisonnable dans le contexte de la gestion du risque de taux. Si une ACP sur base de corrélation est requise, on peut fixer `standardize=True`.

5.2 Vérification du ratio de variance expliquée

L'exécution de l'ACP sur la courbe JGB donne typiquement les ratios de variance expliquée suivants :

Composante principale	Nom	Ratio de variance expliquée
PC1	Niveau (Level)	≈ 70–85%
PC2	Pente (Slope)	≈ 8–15%
PC3	Courbure (Curvature)	≈ 3–8%
Cumulé (3 premières)		≈ 95–99%

Point clé

Plus de 95% de la variation quotidienne de la courbe des taux JGB peut être expliquée par seulement trois facteurs. C'est la base empirique du « modèle à trois facteurs » (Litterman–Scheinkman, 1991), également bien connu sur le marché des Treasuries américains et d'autres marchés obligataires souverains développés.

5.3 Scree plot

La figure 5.1 montre le scree plot. Le panneau de gauche présente le ratio de variance expliquée de chaque composante principale sous forme de diagramme en barres, et le panneau de droite présente le ratio de variance expliquée cumulé sous forme de courbe.

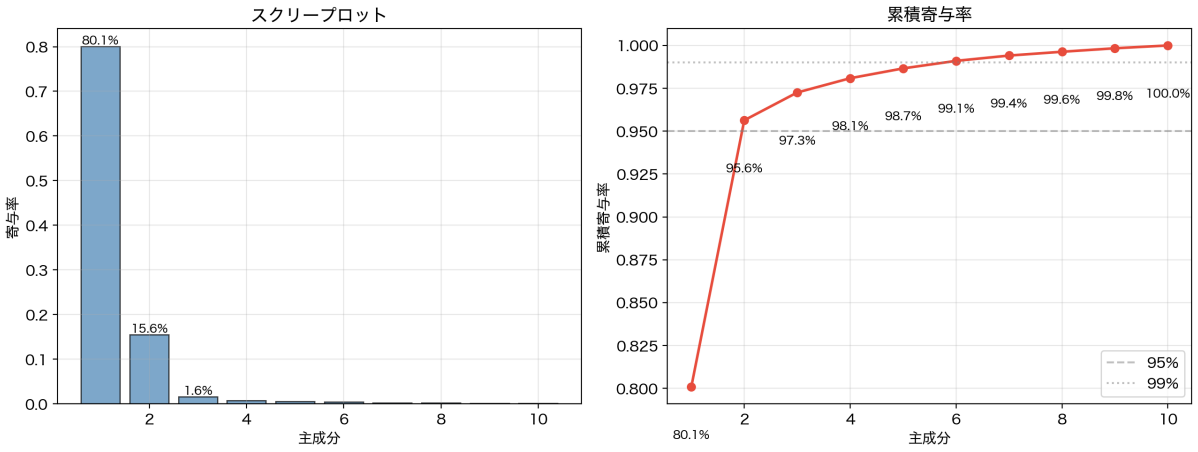

FIGURE 5.1 – Scree plot (à gauche : EVR par composante, à droite : EVR cumulé).

— La décroissance brutale (le « coude ») du diagramme en barres donne un repère pour le nombre de facteurs à retenir.
— Le point où l'EVR cumulé dépasse 95% (ou 99%) constitue également un critère de décision.
— Pour la courbe JGB, $r = 3$ est généralement le nombre de facteurs approprié.

5.4 Série temporelle des scores de facteurs

La figure 5.2 montre la série temporelle des scores des trois composantes principales. PC1 (le facteur de niveau) varie fortement aux points d'inflexion de la politique monétaire, tandis que PC2 (le facteur de pente) tend à être influencé par le cycle économique. PC3 (le facteur de courbure) est relativement bruité mais sert de source de signal pour le trading butterfly.

FIGURE 5.2 – Série temporelle des scores de facteurs de l'ACP.

Chapitre 6

Interprétation économique

6.1 Normalisation du signe

Comme les vecteurs de chargement de l'ACP ont un signe indéterminé (à la fois $\boldsymbol{v}_k$ et $-\boldsymbol{v}_k$ sont des vecteurs propres), les signes doivent être normalisés afin d'attribuer une interprétation économique.

Algorithme de normalisation du signe

Pour chaque composante principale k :

1. Vérifier le signe du chargement $v_k^{(10y)}$ correspondant au ténor 10 ans.
2. Si $v_k^{(10y)} < 0$, poser $\boldsymbol{v}_k \leftarrow -\boldsymbol{v}_k$, $\boldsymbol{f}_k \leftarrow -\boldsymbol{f}_k$.

Ici $\boldsymbol{f}_k$ est le vecteur de scores correspondant. La raison d'utiliser le 10 ans comme ancrage est qu'il s'agit de la maturité de référence du marché JGB, et qu'il est empiriquement garanti que les chargements de tous les facteurs y sont non nuls.

Code 6.1 – Implémentation de la normalisation du signe

```
def normalize_component_sign(loadings, scores, anchor_tenor='10y'):
    loadings_norm = loadings.copy()
    scores_norm = scores.copy()
    for pc in loadings.index:
        if loadings.loc[pc, anchor_tenor] < 0:
            loadings_norm.loc[pc] *= -1
            scores_norm[pc] *= -1
    return loadings_norm, scores_norm
```

6.2 Forme des chargements

La figure 6.1 montre les chargements après normalisation du signe.

6.2.1 PC1 : Niveau (Level)

- Les chargements prennent une **valeur positive quasi constante** sur tous les ténors.
- Un score PC1 positif ⇒ l'ensemble de la courbe des taux se déplace **vers le haut** (hausse des taux).
- Son ratio de variance expliquée de 70–85% est largement dominant.
- Sur le plan macroéconomique, il reflète l'orientation de la politique monétaire, les anticipations d'inflation, etc.

6.2.2 PC2 : Pente (Slope)

— Les chargements sont **négatifs** aux ténors courts et **positifs** aux ténors longs (ou inversement).
— Un score PC2 positif ⇒ la courbe se **pentifie** (l'écart long–court s'élargit).
— Son ratio de variance expliquée est d'environ 8–15%.
— Il tend à refléter les variations d'anticipations du cycle économique.

6.2.3 PC3 : Courbure (Curvature)

— Les chargements sont importants au **ventre** (maturités intermédiaires) et de signe opposé aux **ailes** (maturités courtes et très longues).
— Il prend une forme en « papillon » (butterfly).
— Un score PC3 positif ⇒ le ventre devient **bon marché** (le rendement du ventre monte relativement).
— Son ratio de variance expliquée est de 3–8%, mais c'est une composante importante directement liée au trading de valeur relative.

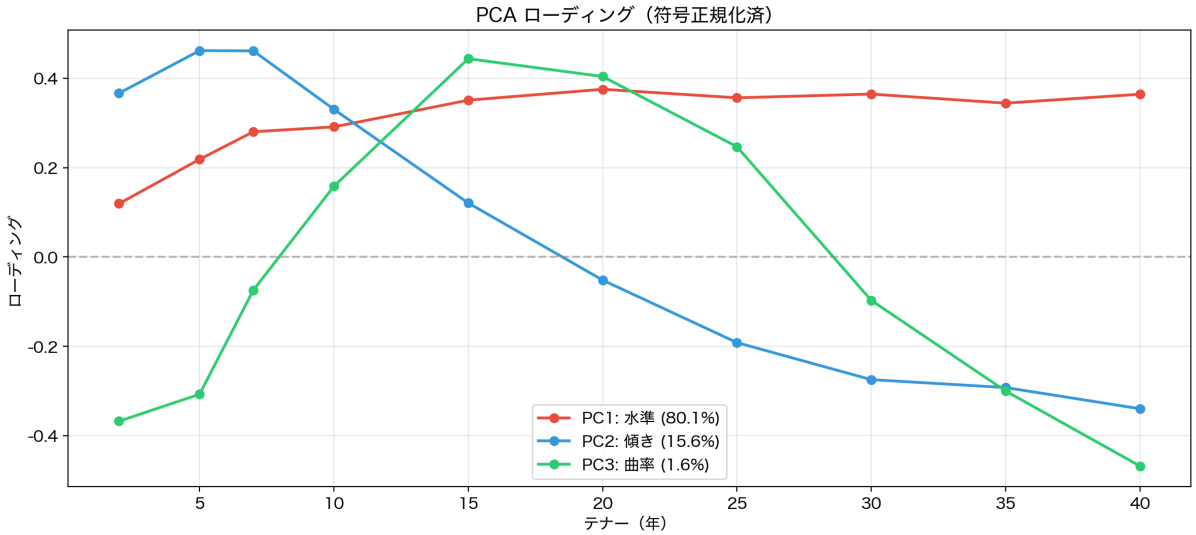

FIGURE 6.1 – Chargements de l'ACP (signe normalisé).

6.3 Décomposition de la contribution des facteurs (reconstruction)

Le vecteur de variation de rendement $\Delta \boldsymbol{y}_t \in \mathbb{R}^{10}$ d'un jour donné t peut être approximativement décomposé par les composantes principales sous la forme

$$\Delta \boldsymbol{y}_t \approx \sum_{k=1}^{r} f_{t,k}\, \boldsymbol{v}_k = f_{t,1}\, \boldsymbol{v}_1 + f_{t,2}\, \boldsymbol{v}_2 + f_{t,3}\, \boldsymbol{v}_3. \tag{6.1}$$

Ici $f_{t,k}$ est le score de la k-ième composante principale au jour t, et $f_{t,k}\, \boldsymbol{v}_k$ représente la « contribution de PCk ».

La figure 6.2 montre la décomposition de la contribution des facteurs pour le jour à plus forte variation des données. La somme des contributions des trois facteurs approche bien les valeurs observées, et le résidu est faible.

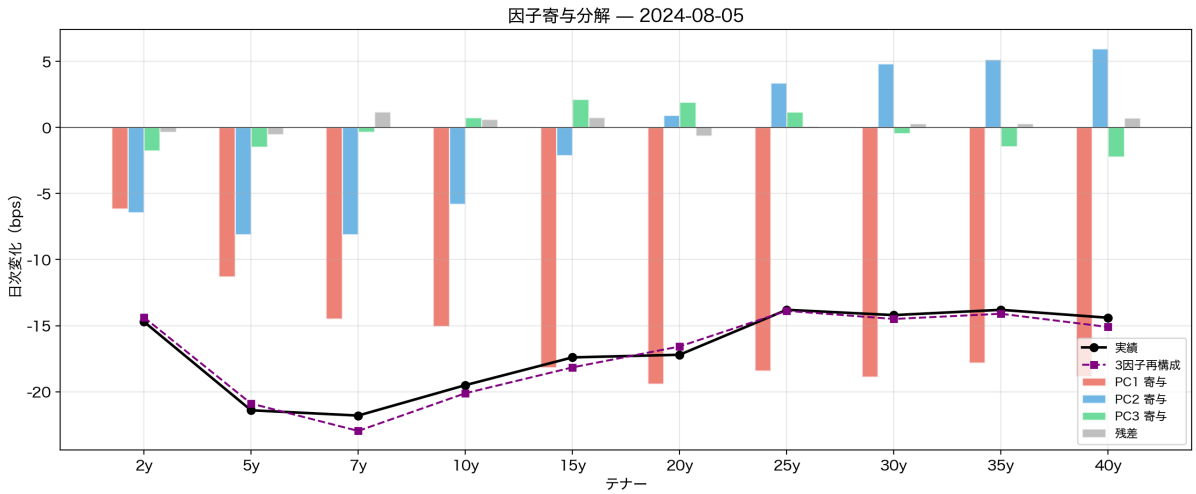

FIGURE 6.2 – Décomposition de la contribution des facteurs : observé (points noirs) vs. reconstruction à 3 facteurs (carrés violets) le jour de forte variation.

Point clé

En effectuant la décomposition de l'équation (6.1) pour un jour où une forte variation s'est produite, on peut déterminer quantitativement si ce mouvement était « dominé par un décalage de niveau », « dominé par une pentification » ou « dominé par une variation de courbure ». Si le résidu (la part non expliquée par les r composantes) est faible, la validité du modèle à trois facteurs est confirmée.

Chapitre 7

De l'ACP au trading

7.1 Transactions butterfly neutres aux facteurs

Une transaction butterfly est composée de trois jambes de ténor (aile courte, ventre, aile longue). Le P&L quotidien d'un butterfly de vecteur de poids $\mathbf{w} = [w_L, w_B, w_R]^\top$ est approché par

$$\mathrm{PnL}_t = \mathbf{w}^\top \Delta \boldsymbol{y}_t^{(\mathrm{legs})} = w_L\, \Delta y_t^{(L)} + w_B\, \Delta y_t^{(B)} + w_R\, \Delta y_t^{(R)} \tag{7.1}$$

(un proxy linéaire).

Définition 7.1 (Butterfly neutre aux facteurs)**.** Lorsque le poids $\mathbf{w}$ satisfait simultanément les conditions suivantes, le butterfly est dit **neutre aux facteurs** (Factor-Neutral) :

$$\text{Neutre en cash :} \quad \mathbf{1}^\top \mathbf{w} = 0 \tag{7.2}$$

$$\text{Neutre PC1 :} \quad \boldsymbol{v}_1^{(\mathrm{legs})\top} \mathbf{w} = 0 \tag{7.3}$$

$$\text{Neutre PC2 :} \quad \boldsymbol{v}_2^{(\mathrm{legs})\top} \mathbf{w} = 0 \tag{7.4}$$

Ici $\boldsymbol{v}_k^{(\mathrm{legs})}$ est la sous-composante à 3 jambes du vecteur de chargement $\boldsymbol{v}_k$.

Cela supprime la contribution au P&L des mouvements de niveau (PC1) et de pente (PC2), n'extrayant que l'**exposition à la courbure pure (PC3)**.

7.2 Solution de norme minimale par SVD

Les conditions sont rassemblées dans une matrice de contraintes $\boldsymbol{C}$:

$$\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 1 \\ v_1^{(L)} & v_1^{(B)} & v_1^{(R)} \\ v_2^{(L)} & v_2^{(B)} & v_2^{(R)} \end{pmatrix} \in \mathbb{R}^{3\times 3}, \qquad \boldsymbol{C}\mathbf{w} = \mathbf{0}. \tag{7.5}$$

$\boldsymbol{C}$ est une matrice 3×3 ; si son rang est 2, l'espace des solutions (noyau) est de dimension 1. Nous déterminons ce noyau par SVD :

1. Décomposer $\boldsymbol{C} = \mathbf{U}_C \mathbf{S}_C \mathbf{V}_C^\top$ par SVD.
2. La dernière colonne de $\mathbf{V}_C$ (correspondant à la plus petite valeur singulière) donne la base $\mathbf{w}^*$ du noyau.
3. Normaliser de sorte que la valeur absolue du poids du ventre soit égale à 1 : $\mathbf{w} = \mathbf{w}^* / |w_B^*|$.
4. Ajuster le signe pour que le ventre soit positif (long).

Code 7.1 – Calcul par SVD des poids butterfly neutres aux facteurs

```python
def compute_fly_weights_svd(loadings_3xk, n_neutral=2):
    C = np.vstack([
        np.ones(3),                          # cash neutral
        loadings_3xk[:, :n_neutral].T        # factor neutral
    ])
    U, S, Vt = np.linalg.svd(C)
    null_dim = Vt.shape[0] - len(S[S > 1e-10])
    w = Vt[-1, :] if null_dim == 0 else Vt[-null_dim:, :].sum(axis=0)
    w = w / np.abs(w[1])    # belly weight = +/- 1
    if w[1] < 0: w = -w     # convention: long belly
    return w
```

La figure 7.1 compare les poids neutres aux facteurs obtenus par SVD pour le fly 5s10s20s aux poids naïfs $[-1, 2, -1]$. Avec les poids neutres aux facteurs par SVD, les sensibilités à PC1 et PC2 sont quasi nulles ($\sim 10^{-6}$), alors qu'avec les poids naïfs des expositions résiduelles à ces facteurs subsistent.

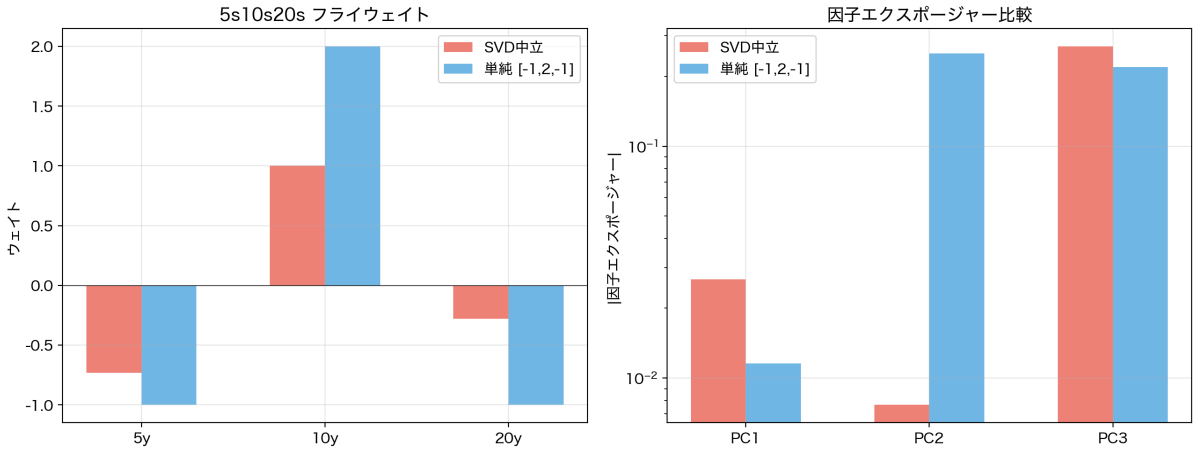

FIGURE 7.1 – Comparaison des poids du fly 5s10s20s (à gauche) et comparaison des expositions aux facteurs (à droite, échelle logarithmique).

Point clé

Grâce à la neutralisation des facteurs, la contribution au P&L des mouvements de niveau (PC1) et de pente (PC2) est supprimée, et l'**exposition à la courbure pure (PC3)** peut être extraite. Avec les poids naïfs $[-1, 2, -1]$, des expositions à PC1 et PC2 d'un ordre de grandeur supérieur subsistent.

Chapitre 8

Backtest walk-forward

8.1 Principes de conception

La chose la plus importante dans un backtest est d'éliminer complètement le **biais d'anticipation** (référence à des données futures, *look-ahead bias*). Ce backtest est conçu comme suit.

Élément	Réglage
Fenêtre d'apprentissage	126 jours ouvrés (glissante)
Pas de test	1 jour (réajustement quotidien)
Signal	Z-score de l'écart du fly
Condition d'entrée	$\lvert z\rvert > z_{\text{in}} = 1.5$
Condition de sortie	$\lvert z\rvert < z_{\text{out}} = 0.5$ ou durée de détention max dépassée
Durée de détention max	20 jours ouvrés
Coûts de transaction	0.2 bps aller simple (prélevés à l'entrée/sortie)

8.2 Mécanisme d'élimination du biais

Backtest walk-forward

Pour chaque jour $t = T_w, T_w + 1, \ldots, T$ (où T_w est la taille de la fenêtre d'apprentissage) :

1. Ajuster l'ACP sur les T_w derniers jours de données $\mathbf{X}_{[t-T_w:t]}$.
2. Calculer les poids neutres aux facteurs par SVD $\mathbf{w}_t$ à partir des chargements courants.
3. Calculer le Z-score z_t de l'écart cumulé du fly.
4. Décider la position p_{t+1} à partir de z_t (appliquée le jour suivant).
5. $\text{PnL}_{t+1} = p_t \times (\mathbf{w}_t^\top \Delta \boldsymbol{y}_{t+1}^{(\text{legs})})$.

Attention

— Le signal au temps t n'utilise **que** les données jusqu'au temps t.
— La position est décalée d'un jour et appliquée en $t + 1$.
— Ni signaux ni positions n'existent pour la période antérieure à la fenêtre d'apprentissage.

8.3 Logique de trading

La position prend trois états $p_t \in \{-1, 0, +1\}$:

— $p_t = 0$ (plat) : aucune position.
— $p_t = +1$ (ventre long) : le fly est bon marché ($z < -z_{\text{in}}$).
— $p_t = -1$ (ventre court) : le fly est cher ($z > +z_{\text{in}}$).

La sortie est déclenchée par l'une des conditions suivantes :

— En position ventre long, $z > -z_{\text{out}}$ (le retour à la moyenne a progressé).
— En position ventre court, $z < +z_{\text{out}}$.
— La durée de détention dépasse `max_hold`.

8.4 Traitement des coûts de transaction

Les coûts de transaction prélèvent le montant aller simple à chaque changement de position :

$$\text{Cost}_t = |\Delta p_t| \times c_{\text{one-way}}, \tag{8.1}$$

où $\Delta p_t = p_t - p_{t-1}$ et $c_{\text{one-way}} = 0.2$ bps. Le P&L net est calculé par

$$\text{NetPnL}_t = \text{RawPnL}_t - \text{Cost}_t. \tag{8.2}$$

8.5 Indicateurs de performance

Les résultats du backtest sont évalués à l'aide des indicateurs suivants.

Indicateur	Définition	Signification
Total PnL	$\sum_t \text{NetPnL}_t$	P&L cumulé
Vol. ann.	$\sigma_{\text{daily}} \times \sqrt{252}$	Volatilité annualisée
Ratio de Sharpe	$\dfrac{\mu_{\text{daily}} \times 252}{\sigma_{\text{daily}} \times \sqrt{252}}$	Rendement ajusté du risque
Max Drawdown	$\min_t(\text{CumPnL}_t - \max_{s \le t} \text{CumPnL}_s)$	Perte cumulée maximale
Taux de réussite	$P(\text{PnL} > 0 \mid p_t \neq 0)$	Taux de gain
Détention moy.	jours actifs / nombre d'aller-retours	Durée de détention moyenne

8.6 Résultats du backtest

La figure 8.1 compare le P&L cumulé du fly neutre aux facteurs par SVD et du fly naïf $[-1, 2, -1]$ pour quatre configurations de fly.

La figure 8.2 présente le détail du signal (Z-score), de la position et du P&L cumulé pour le fly au meilleur ratio de Sharpe.

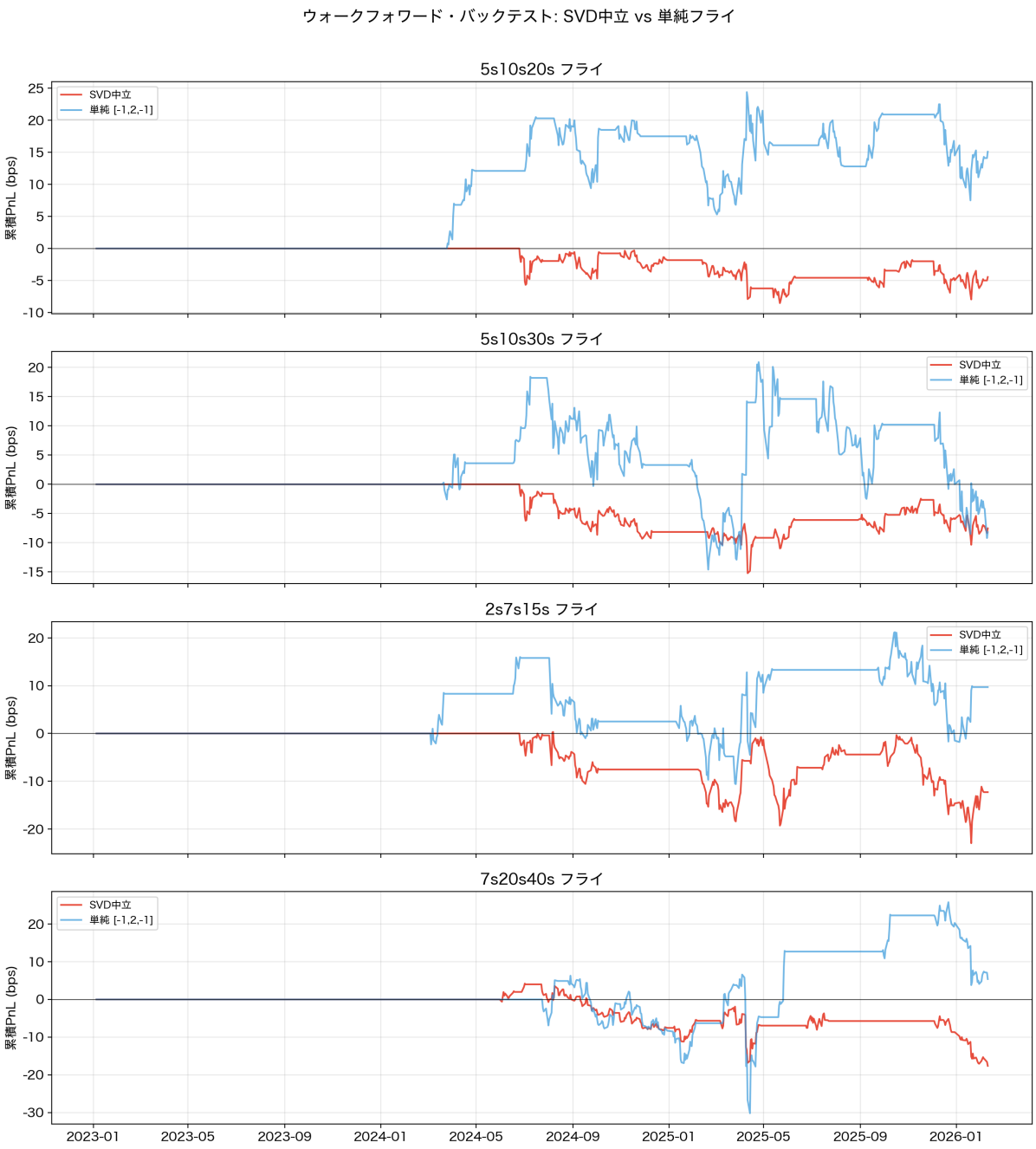

FIGURE 8.1 – Backtest walk-forward : P&L cumulé du fly neutre aux facteurs par SVD vs. fly naïf.

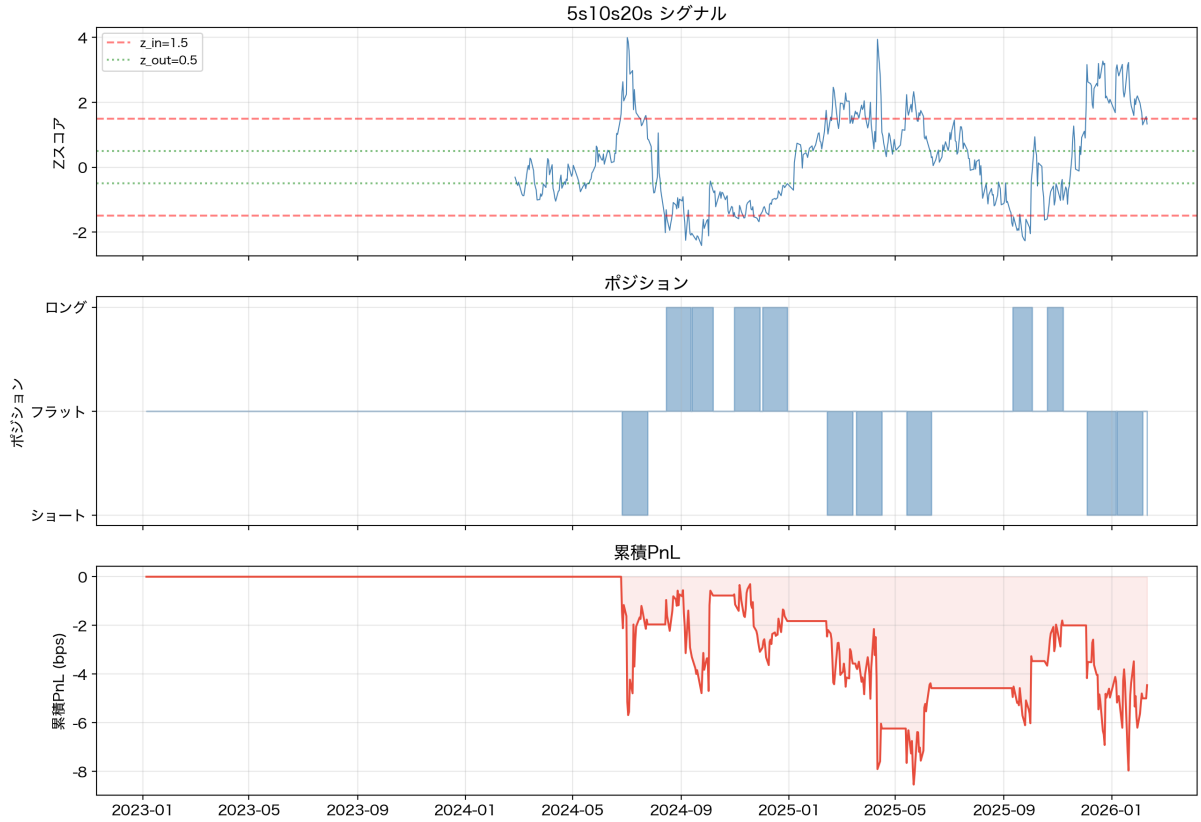

FIGURE 8.2 – Analyse du signal du meilleur fly : Z-score (haut), position (milieu), P&L cumulé (bas).

Chapitre 9

Validation de la robustesse

Pour vérifier la robustesse de la stratégie, nous effectuons une **analyse de sensibilité** aux principaux paramètres.

9.1 Sensibilité à la fenêtre d'apprentissage

Nous faisons varier la fenêtre d'apprentissage sur $\{63, 126, 252\}$ jours ouvrés (environ 3 mois, 6 mois, 1 an) et vérifions l'impact sur la performance (figure 9.1).

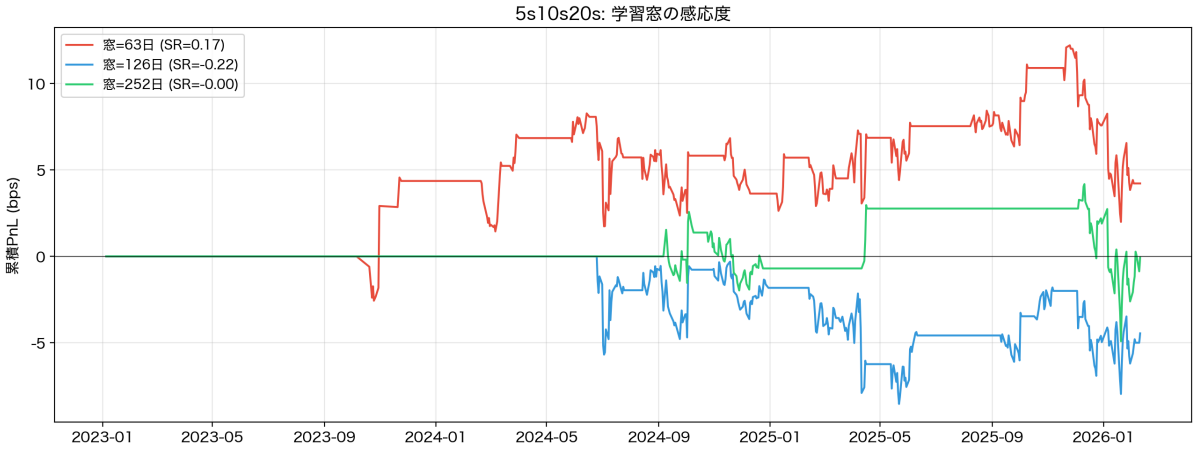

FIGURE 9.1 – Sensibilité à la taille de la fenêtre d'apprentissage.

— **Fenêtre courte** (63 jours) : s'adapte rapidement à la structure de courbe la plus récente mais est sensible au bruit.
— **Fenêtre moyenne** (126 jours) : un bon équilibre entre stabilité et adaptativité.
— **Fenêtre longue** (252 jours) : fournit des estimations stables mais réagit lentement aux changements de régime.

9.2 Sensibilité aux seuils Z

Nous validons des combinaisons de seuils d'entrée/sortie par une recherche sur grille :

z_{in}	z_{out}	Caractéristiques
1.0	0.3	Trading haute fréquence, forte charge de coûts
1.5	0.5	Référence
2.0	0.5	Entrée conservatrice
2.0	1.0	Courte durée de détention, forte rotation

En général, des seuils plus bas augmentent la fréquence de trading et la charge de coûts, tandis que des seuils plus élevés réduisent les opportunités de trading mais améliorent la précision du signal.

9.3 Sensibilité au nombre de facteurs

Nous comparons la neutralisation de 2 facteurs (PC1 et PC2 seulement) et de 3 facteurs (PC1–PC3).

— **Neutre à 2 facteurs :** outre l'exposition à la courbure (PC3), une exposition aux résidus d'ordre supérieur subsiste également. C'est l'intention de la stratégie.

— **Neutre à 3 facteurs :** la position devient neutre à toutes les composantes PC1–PC3, mais imposer trois contraintes (dont la neutralité en cash) à trois jambes rend la matrice de contraintes $\boldsymbol{C}$ quasi régulière, ne laissant presque aucun degré de liberté. En conséquence, les poids tendent à devenir instables.

Chapitre 10

Estimation de la fenêtre optimale et stabilité des composantes

Dans une ACP glissante, la taille de la fenêtre d'apprentissage est un hyperparamètre important qui détermine la qualité d'estimation et l'adaptativité du modèle. De plus, avec les changements de taille de fenêtre ou de régime de marché, les chargements (vecteurs propres) des composantes principales varient dans le temps, et l'interprétation économique (niveau / pente / courbure) peut se rompre.

Dans ce chapitre, nous discutons (1) d'une méthode systématique d'estimation de la taille de fenêtre optimale, (2) d'une évaluation quantitative de la variation temporelle des chargements, et (3) de contre-mesures pratiques pour préserver l'interprétabilité.

10.1 Estimation de la fenêtre optimale

10.1.1 Compromis biais–variance

Le choix de la taille de fenêtre T_w implique le compromis suivant :

— **Si la fenêtre est trop courte** ($T_w \ll 100$) : le nombre d'échantillons est faible et l'erreur d'estimation (variance) de la matrice de covariance est grande. Les chargements deviennent bruités et les poids neutres aux facteurs deviennent instables.

— **Si la fenêtre est trop longue** ($T_w \gg 200$) : une structure ancienne contamine l'estimation, et la structure actuelle de co-mouvement de la courbe ne peut pas être correctement reflétée (le biais est grand). En particulier, après un changement de régime, les données de l'ancien régime faussent l'estimation.

Pour évaluer quantitativement ce compromis biais–variance, nous utilisons l'erreur de reconstruction hors échantillon (OOS).

10.1.2 Erreur de reconstruction OOS

Nous définissons la RMSE de reconstruction OOS pour une taille de fenêtre T_w comme suit.

Définition 10.1 (Erreur de reconstruction OOS)**.** Pour chaque jour $t > T_w$, estimer une ACP à r facteurs sur les T_w jours de données immédiatement précédents et reconstruire le vecteur de variation de rendement $\Delta \boldsymbol{y}_t$ du jour t. L'erreur de reconstruction est

$$e_t(T_w) = \left\| \Delta \boldsymbol{y}_t - \hat{\mathbf{V}}_r(t)\, \hat{\mathbf{V}}_r(t)^\top \Delta \boldsymbol{y}_t \right\|_2, \tag{10.1}$$

où $\hat{\mathbf{V}}_r(t) \in \mathbb{R}^{p \times r}$ est la matrice des chargements des r premières composantes estimée à partir des données $\mathbf{X}_{[t-T_w:t-1]}$.

La RMSE OOS est calculée par

$$\mathrm{RMSE}(T_w) = \sqrt{\frac{1}{T - T_w} \sum_{t=T_w+1}^{T} e_t(T_w)^2}. \tag{10.2}$$

Remarque 10.1. Dans l'équation (10.1), $\hat{\mathbf{V}}_r(t)\,\hat{\mathbf{V}}_r(t)^\top$ est la matrice de projection orthogonale sur le sous-espace de dimension r. Le résidu de la reconstruction par cette projection constitue l'erreur OOS. Cela mesure dans quelle mesure le sous-espace appris sur les données de la fenêtre peut expliquer les données du jour suivant.

La figure 10.1 montre la RMSE OOS pour chaque taille de fenêtre $T_w = 30, 40, \ldots, 300$.

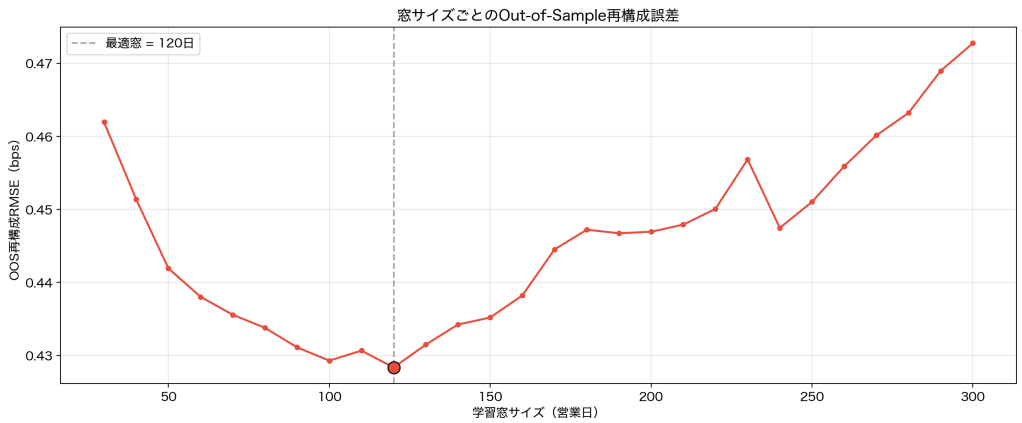

FIGURE 10.1 – RMSE de reconstruction OOS pour chaque taille de fenêtre. La taille de fenêtre donnant le minimum est l'estimation optimale.

Point clé

La taille de fenêtre qui minimise la RMSE OOS correspond au point où biais et variance s'équilibrent. Pour des fenêtres trop courtes, l'erreur due à la variance domine, tandis que pour des fenêtres trop longues, le biais dû au changement structurel domine. Ce point optimal dépend du degré de non-stationnarité des données, et peut donc varier avec l'environnement de marché.

10.1.3 Stabilité de l'EVR

Un autre indicateur d'évaluation de la taille de fenêtre est la **stabilité temporelle** du ratio de variance expliquée (EVR). En exécutant une ACP glissante avec une taille de fenêtre T_w, nous calculons l'écart-type de l'EVR de chaque PC :

$$\sigma_{\mathrm{EVR}_k}(T_w) = \mathrm{Std}\left[\mathrm{EVR}_k(t; T_w)\right]_{t=T_w+1}^{T}. \tag{10.3}$$

Plus σ_{EVR_k} est petit, plus le pouvoir explicatif de la k-ième composante principale est stable à cette taille de fenêtre.

La figure 10.2 montre l'écart-type de l'EVR par taille de fenêtre (à gauche) et la série temporelle de l'EVR glissant (à droite).

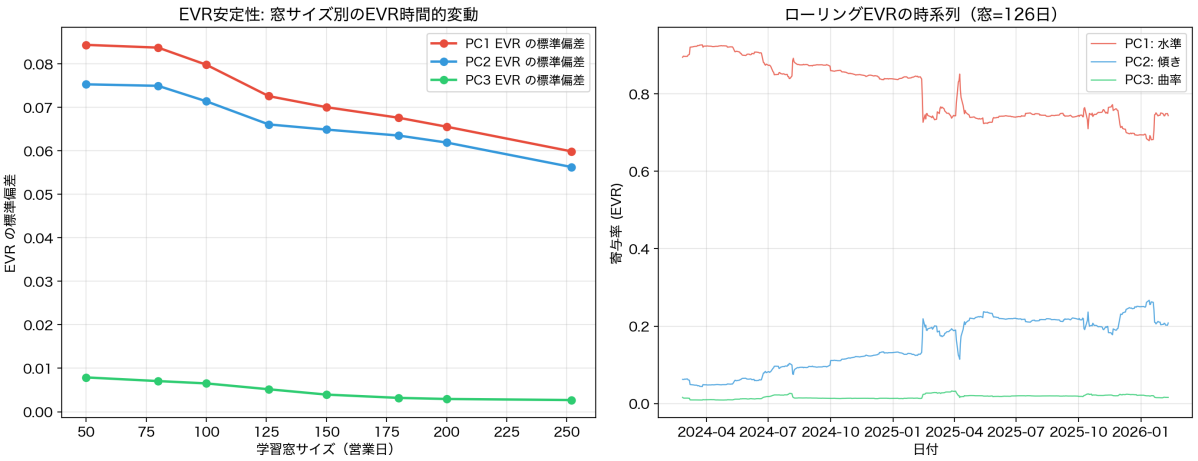

FIGURE 10.2 – Stabilité de l'EVR : variation de l'EVR par taille de fenêtre (à gauche) et série temporelle de l'EVR glissant (à droite, fenêtre = 126 jours).

10.2 Variation temporelle des chargements

10.2.1 Évaluation quantitative par similarité cosinus

Nous mesurons à quel point le chargement $\boldsymbol{v}_k(t)$ de l'ACP glissante (la k-ième direction principale au temps t) s'écarte du chargement de la période complète $\boldsymbol{v}_k^{(\text{full})}$ à l'aide de la **similarité cosinus**.

Définition 10.2 (Similarité cosinus des chargements)**.**

$$\cos\theta_k(t) = \frac{\left|\boldsymbol{v}_k(t)^\top \boldsymbol{v}_k^{(\text{full})}\right|}{\|\boldsymbol{v}_k(t)\|\,\|\boldsymbol{v}_k^{(\text{full})}\|}. \tag{10.4}$$

Prendre la valeur absolue ici sert à absorber l'indétermination du signe des vecteurs propres. $\cos\theta_k(t) = 1$ signifie une concordance parfaite, et $\cos\theta_k(t) \approx 0$ signifie l'orthogonalité (une direction complètement différente).

Remarque 10.2. Un instant où la similarité cosinus chute brusquement suggère un candidat de rupture structurelle (*structural break*) dans l'ACP. À de tels instants, le recalcul des poids neutres aux facteurs devient particulièrement important.

La figure 10.3 montre la série temporelle de la similarité cosinus pour PC1–PC3.

— **PC1 (Niveau)** maintient $\cos\theta \approx 1$ sur presque toute la période. La structure dans laquelle tous les ténors évoluent dans la même direction est très stable.

— **PC2 (Pente)** a généralement $\cos\theta > 0.9$, mais une baisse temporaire peut être observée aux points d'inflexion de la politique monétaire.

— **PC3 (Courbure)** est relativement instable, avec des périodes où $\cos\theta$ chute en dessous de 0.8. Cela reflète le fait que PC3 a un faible ratio de variance expliquée et est facilement affecté par le bruit, en plus du « basculement » possible entre PC2 et PC3.

10.2.2 Carte de chaleur de la variation temporelle des chargements

La figure 10.4 visualise la variation temporelle des chargements PC1–PC3 sous forme de carte de chaleur. L'axe horizontal est la date, l'axe vertical est le ténor, et la couleur représente la valeur du chargement.

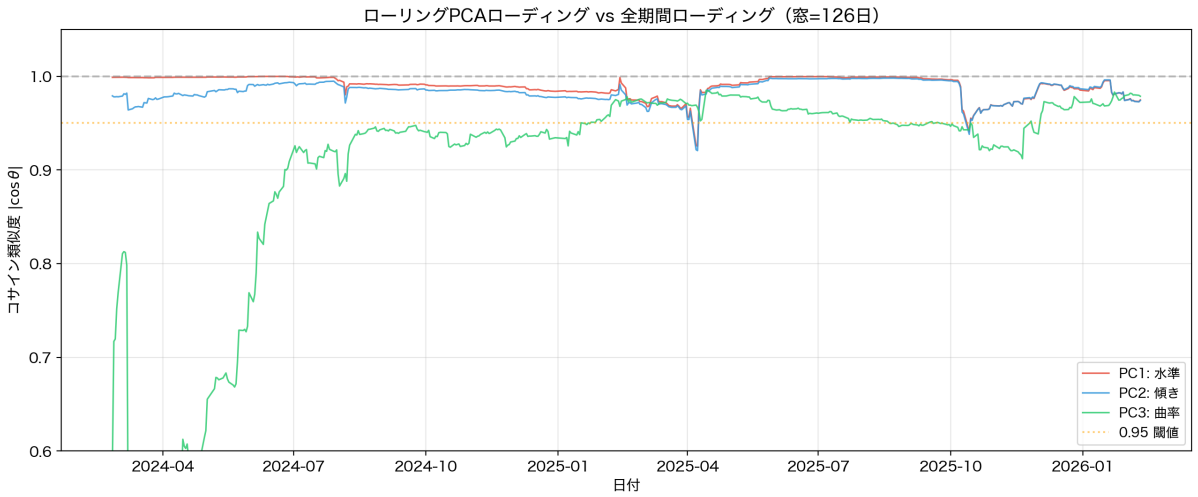

FIGURE 10.3 – Similarité cosinus des chargements de l'ACP glissante (fenêtre = 126 jours, comparée aux chargements de la période complète).

On peut confirmer que, tandis que le chargement de PC1 maintient de façon stable des valeurs positives sur tous les ténors, la « position du pic » du chargement de PC3 le long de l'axe des ténors varie dans le temps. Une telle variation est l'une des causes de l'instabilité des poids du fly.

10.3 Préservation de l'interprétabilité et mélange des composantes

10.3.1 Basculement des PC

Dans une ACP glissante, pendant les périodes où les valeurs propres de PC2 et PC3 sont proches, un « basculement » (croisement de valeurs propres, *eigenvalue crossing*) des chargements correspondants peut survenir. À ce moment, une inversion temporaire se produit dans laquelle PC2 prend une forme de type courbure et PC3 une forme de type pente.

Ce basculement peut être détecté de la manière suivante :

1. Lorsque le ratio d'EVR de PC adjacentes, λ_k/λ_{k+1}, approche de 1, le risque de basculement est élevé.
2. En termes de similarité cosinus, le basculement s'est produit lorsque PC2 est proche de la direction de PC3 de la période complète et que PC3 est proche de la direction de PC2 de la période complète.

10.3.2 Contre-mesures pratiques

Nous résumons les principales contre-mesures contre la variation temporelle des chargements et la rupture de l'interprétabilité.

Alignement sur des chargements de référence. Réordonner les chargements obtenus par l'ACP glissante de sorte que leur produit scalaire avec les chargements de référence de la période complète (ou de la dernière année) soit maximisé. Cela corrige automatiquement le basculement PC2/PC3.

Rotation de Procrustes. Pour une matrice de chargements de référence $\mathbf{V}^{(\mathrm{ref})}$, transformer les chargements glissants $\mathbf{V}(t)$ par une rotation orthogonale $\boldsymbol{Q}^*$ pour les rapprocher au maximum :

$$\boldsymbol{Q}^* = \arg\min_{\boldsymbol{Q}\boldsymbol{Q}^\top = \boldsymbol{I}} \left\| \mathbf{V}(t)\boldsymbol{Q} - \mathbf{V}^{(\mathrm{ref})} \right\|_F . \tag{10.5}$$

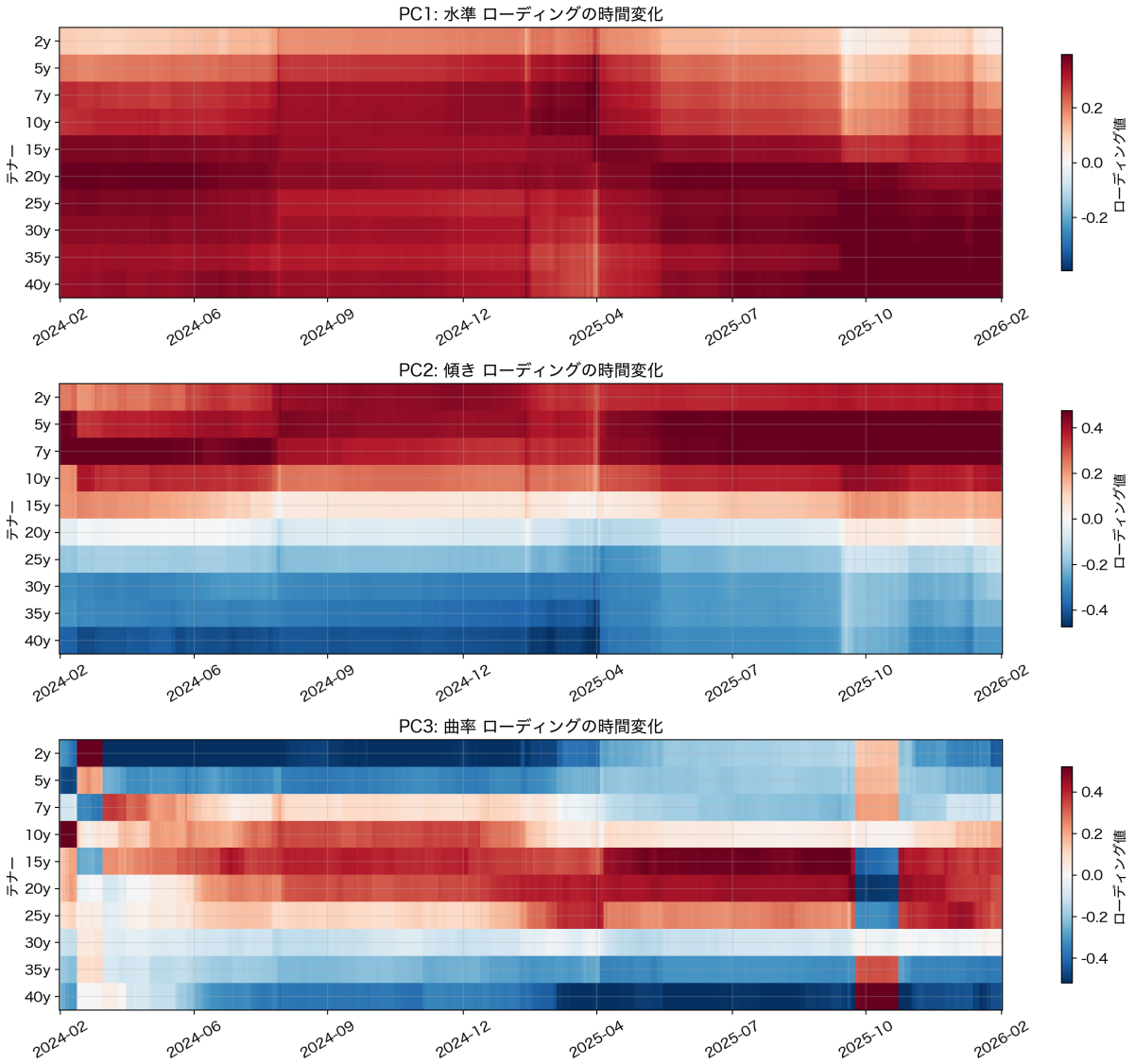

FIGURE 10.4 – Carte de chaleur de la variation temporelle des chargements (fenêtre = 126 jours). De haut en bas : PC1 (Niveau), PC2 (Pente), PC3 (Courbure).

Ceci s'obtient sous forme close $\boldsymbol{Q}^* = \mathbf{V}'\mathbf{U}'^\top$ à partir de la SVD $\mathbf{V}^{(\mathrm{ref})\top}\mathbf{V}(t) = \mathbf{U}'\mathbf{S}'\mathbf{V}'^\top$. La rotation de Procrustes maintient l'interprétation économique des chargements cohérente dans le temps.

Taille de fenêtre adaptative. Surveiller la stabilité de l'EVR et la similarité cosinus, et raccourcir la taille de fenêtre lorsqu'un changement structurel est détecté. En période stable, allonger la fenêtre pour améliorer la précision d'estimation ; en période volatile, raccourcir la fenêtre pour améliorer l'adaptativité.

Attention

Si les poids neutres aux facteurs sont calculés dans un état où l'interprétabilité des chargements s'est rompue, des expositions involontaires aux facteurs apparaissent. En particulier, durant les périodes où un basculement PC2/PC3 se produit, une situation peut survenir où ce qui est censé être « neutre à la pente » est devenu « neutre à la courbure ». Il est recommandé de valider séparément la performance de telles périodes dans le backtest.

Chapitre 11

Mises en garde et extensions

11.1 Mises en garde

1. **Effets du roll du CTD :** le titre CTD des contrats à terme JGB se situe souvent dans la zone 7 ans, et des sauts peuvent apparaître dans les rendements à maturité constante lors des rollovers de contrats. De tels sauts peuvent générer des signaux d'ACP fallacieux. Une attention particulière est requise pour les configurations de fly incluant le ténor 7 ans.
2. **Biais du survivant :** comme une configuration de fly prometteuse est sélectionnée ex post, la sélection de la configuration dans l'échantillon introduit un biais. Pour une véritable validation OOS, la sélection de la configuration de fly elle-même doit également être fixée à l'avance.
3. **Éviter le surajustement :** la conception walk-forward atténue le risque de surajustement mais ne l'élimine pas complètement. Il est important de ne pas calibrer les paramètres sur des données ultérieures une fois qu'ils ont été choisis.
4. **Proxy linéaire du P&L :** le P&L est une approximation linéaire des variations de rendement, tandis que le P&L réel inclut la duration, la convexité et les effets de portage/roll-down. En particulier, dans la zone des maturités longues, les effets de convexité peuvent ne pas être négligeables.
5. **Coûts de transaction :** le chiffre de 0.2 bps aller simple est une valeur hypothétique. Les coûts réels dépendent de l'environnement de liquidité du marché, de l'écart bid–offer, de la taille de lot de transaction, etc. En période de stress, les écarts peuvent s'élargir considérablement.

11.2 Sujets avancés

ACP sur ASW (ASW-PCA). Appliquer l'ACP aux variations des spreads d'asset-swap pour construire des signaux de valeur relative. Comme les spreads ASW reflètent la valeur relative nette des coûts de financement, ils peuvent servir de signal de trading plus direct en pratique.

ACP pondérée par la duration. Pondérer les variations par ténor par le DV01 pour construire des facteurs cohérents avec le risque. Cela rend la neutralisation des facteurs exacte sur une base de risque et améliore également la précision de l'approximation du P&L.

Détection de régime. Détecter les régimes de volatilité de PC1 et construire des stratégies conditionnelles, comme relever le seuil d'entrée dans les environnements de forte volatilité. Des

combinaisons avec des modèles à changement de régime markovien ou des modèles GARCH peuvent être envisagées.

Neutralisation multi-jambes. Étendre le butterfly à 3 jambes à 4 jambes ou plus et implémenter des versions neutres à la duration ou au DV01. Augmenter le nombre de jambes ajoute des degrés de liberté aux contraintes et permet une neutralisation des facteurs plus précise.

Liste des symboles

Symbole	Signification
$\mathbf{X} \in \mathbb{R}^{n \times p}$	Matrice de données centrée (n : échantillons, p : variables)
$\boldsymbol{\Sigma}$	Matrice de covariance empirique
λ_k	k-ième valeur propre
$\boldsymbol{v}_k$	k-ième direction principale (chargement)
$f_{t,k}$	Score de la k-ième composante principale au jour t
EVR_k	Ratio de variance expliquée de la k-ième PC
$\Delta y_t^{(k)}$	Variation quotidienne du rendement du ténor k (bps)
$\mathbf{w}$	Vecteur de poids du butterfly
$\boldsymbol{C}$	Matrice de contraintes (neutre en cash + neutre aux facteurs)
z_t	Z-score de l'écart du fly
T_w	Taille de la fenêtre d'apprentissage (jours ouvrés)
$z_{\text{in}}, z_{\text{out}}$	Seuils d'entrée / sortie